NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

MULTIVARIATE ANALYSIS OF THE EFFECT OF GRADUATE EDUCATION ON PROMOTION TO ARMY LIEUTENANT COLONEL

by

Hakan Kabalar

June 2003

Thesis Advisor: Samuel E.Buttrey
Co-Advisor: Robert A.Koyak

Approved for public release; distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY	2. REPORT DATE June 2003	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE Multivariate Analysis of the Effect of Graduate Education on Promotion to Army Lieutenant Colonel	5. FUNDING NUMBERS
6. AUTHOR (S) Kabalar, Hakan	

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the U.S. Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

The objective of this thesis is to estimate and explain the effects of graduate education and other factors on promotion to the rank of Lieutenant Colonel (O-5) in the US Army. Our focus was primarily on determining whether graduate education provides officers with higher promotion probabilities. Besides graduate education, data that were analyzed include basic demographic traits, the officers' prior enlisted status, and their commissioning source information. The data used in this study were taken from the Active Duty Military Master File for fiscal years 1981 through 2001.

This study develops multivariate logit regression and classification tree models to examine and explore the structure of the data sets. Both the regression models and the classification trees yielded positive results for the effect of graduate education on promotion. According to the regression model results, the odds ratio associated with graduate education is between 1.79 and 2.25. Military Academy and ROTC/Scholarship graduates have higher promotion probabilities than those from other sources, and married officers have higher rates than single officers. Additionally, age has a negative effect on promotion; that is, promotion probability decreases with age. Prior enlisted status, number of dependents, gender, race, and DOD primary occupation code do not seem to have statistically significant effects on promotion.

14. SUBJECT TERMS Graduate Education, Promotion, Logistic Regression, and Classification Tree.	15. NUMBER OF PAGES 73
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

MULTIVARIATE ANALYSIS OF THE EFFECT OF GRADUATE EDUCATION ON PROMOTION TO ARMY LIEUTENANT COLONEL

Hakan Kabalar
First Lieutenant, Turkish Army
B.S., Turkish Army Academy, 1998

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

**NAVAL POSTGRADUATE SCHOOL
June 2003**

Author:

Hakan Kabalar

Approved by:

Samuel E.Buttrey
Thesis Advisor

Robert A. Koyak
Second Reader

James N. Eagle
Chairman, Department of Operations Research

iii

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The objective of this thesis is to estimate and explain the effects of graduate education and other factors on promotion to the rank of Lieutenant Colonel (O-5) in the US Army. Our focus was primarily on determining whether graduate education provides officers with higher promotion probabilities. Besides graduate education, data that were analyzed include basic demographic traits, the officers' prior enlisted status, and their commissioning source information. The data used in this study were taken from the Active Duty Military Master File for fiscal years 1981 through 2001.

This study develops multivariate logit regression and classification tree models to examine and explore the structure of the data sets. Both the regression models and the classification trees yielded positive results for the effect of graduate education on promotion. According to the regression model results, the odds ratio associated with graduate education is between 1.79 and 2.25. Military Academy and ROTC/Scholarship graduates have higher promotion probabilities than those from other sources, and married officers have higher rates than single officers. Additionally, age has a negative effect on promotion; that is, promotion probability decreases with age. Prior enlisted status, number of dependents, gender, race, and DOD primary occupation code do not seem to have statistically significant effects on promotion.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Graduate education is deemed a necessary tool to enhance leaders' effectiveness to meet the U.S. Army's needs in an era of burgeoning technological developments, uncertain battlegrounds, unconventional military strategies, and sophisticated weapon systems. Many officers, however, see the approximately two-year term spent for graduate education at schools apart from their principal military services as a setback for their future progress in the Army. This perception was the primary reason for the study.

The main purpose of this study is to examine the impact of graduate education on promotion to the rank of Lieutenant Colonel (O-5) in the U.S. Army. Despite the fact that there have been a number of studies on the effect of graduate education on promotion and retention of U.S. Navy officers, there have not been many similar studies related to the Army.

The data used in this study were taken from the Active Duty Military Master File for fiscal years 1981 through 2001 as cohort data sets. Combining nine predictor and one response variable 1981, 1982, 1983 and pooled data sets were established. The ten-year-point was selected to represent those variables tracked yearly. Demographic variables consisted of Gender, Race, Age, Marital Status, and Number of Dependents. Professional factors were Education, Commission Source, DOD Primary Occupation Code (DPOG) and Prior Enlisted (PE). The response variable was Promotion to O-5.

Logistic regression and classification tree models were utilized to explain and explore the effect of each predictor variable on promotion. The results indicated that officers with graduate education have higher promotion probabilities. The estimated odds ratio for graduate education was within a range of 1.79 to 2.25 suggesting that the promotion ratio among officers with graduate degrees is 1.79-2.25 times the same ratio for officers without graduate degrees.

Another important outcome was that Academy or ROTC/Scholarship graduates have higher promotion probabilities compared to officers from other commission sources. Being married also increases the odds ratio for promotion. On the other hand, age has a negative effect on promotion; that is, the promotion probability decreases with an increase in age. The number of dependents, DOD primary occupation code, gender, and prior enlisted variables were found not to be statistically significant.

Classification tree results supported the regression model results on the positive impact of graduate education on promotion. Commission Source, Education and Age were the primary variables that had impact on promotion rates. According to the tree models, officers with a graduate degree, who graduated from Academy or ROTC/Scholarship program, and who were 33.5 years old or younger in their tenth year of service have the highest predicted probabilities of promotion.

I.   INTRODUCTION

A.   BACKGROUND

Military operations in the 20th century including from various level of engagements to peacekeeping operations, such as Bosnia, Kosovo, and recent Gulf War operations, signal that we should expect uncertainty on the battleground. There is an increasing need for sophisticated leadership and flexible plans to synchronize personnel, weapon systems, information sources, and technology to outperform potential adversaries. Experiments involving Force XXI and the Army After Next have emphasized the Army's developing capabilities and its associated personnel requirements, with a keen appreciation and understanding of the historical perspective of war:

> The results suggest that officers need to acquire new analytical and cognitive skills for thinking through and solving complex military problems today and in the future. The success of Army units will continue to depend upon leaders who are intellectually agile, logical, creative, and innovative in their thinking, and who are also bold and audacious decision-makers. The nature of future military problems and operating environments may demand new skills, knowledge, and attributes of officers. For example, recent Force XXI Army war-fighting experiments at Fort Hood, Texas, and the National Training Center (NTC), Fort Irwin, California, demonstrated the need for officers who are comfortable commanding, leading, and managing in high-technology environments. Mission success of Force XXI units also depended upon how well the leaders managed and operated Command, Control, Communications, Computers, and Intelligence (C4I) systems to build and maintain a relevant

common picture of the battle space and synchronize military operations. The extent to which we exploit future advances in military and C4I technologies will eventually determine military troops' effectiveness as a fighting force. These developments require officers able to perform essential functions that fall outside of the Army's war-fighting role but are absolutely necessary to field an Army that can fight and win. (Officer Personnel Management System XXI Study, 1997)

This excerpt clarifies the anticipated ambiance of the future battlefield and the need for leadership with technical and specialty skills in addition to battleground experience to live up to the Army's needs. Imagine an Army of the 21st century that has the most sophisticated weapons systems in the world, and that this Army has the most challenging missions that take it to multiple regions of the world within a few hours to fight against any conceivable type of enemy. Imagine further the substantial changes in this century's operating, technological, and economic environments, all of which pose only uncertainty. Given the Army's vision to accomplish strategic dominance across the entire spectrum of operations, factors such as responsiveness, deployability, agility, versatility, lethality, survivability, and sustainability have paramount importance not only in the transformation of the Army but also in meeting the criteria of being a robust and unchallengeable Army.

The Army conducts military training in peacetime as preparation for military operations. Education and training are the only means by which not only to enhance the effectiveness of its preparation for such an uncertain

future but also to improve the ability of personnel to adapt more quickly to the changing environment. In particular, it would not be wrong to assert that graduate education can provide the Army with a more productive and versatile officer corps to meet the needs of the Army by synthesizing military training with academics. A National Academy of Sciences study emphasizes the value of graduate education:

> Graduate education provides career-long enhancement of the abilities of an officer, not just a technical specialty skill. Development of problem-solving skills is applicable to all kinds of problems that face the individual in unexpected situations. It is self-evident that there is little time for such education in wartime. The time to devote resources to obtaining graduate education is when the nation is at peace. It should then be a high priority whose payoff is enhanced performance in times of war as well as in times of peace. Graduate education is a generator of future readiness with a high rate of return. (National Academy of Sciences 1997, Volume 4, p.39)

Although from top leadership to the lowest-ranking military officers the importance of graduate education is recognized, there are mixed perceptions about its role in officers' career development, especially in promotion to crucial military posts on the path of career development.

Besides the obvious advancements in science and technology evident in the Army's war-fighting equipment, the increase in information and the more-detailed decision-making required in modern doctrine and warfare necessitates increased specialization within the officer corps. Complex lethal weapons, joint and multinational

doctrine and organizations, global political and economic connectivity require the utmost technical competence in the officer corps.

B.   OBJECTIVES

The purpose of this thesis is to clarify the importance of graduate education for officers' career development.

This study examined the effects of graduate education and other factors on promotion to the rank of Army Lieutenant Colonel (O-5). The emphasis was on the following question: "Is there a statistically significant difference in the rate of promotion to the rank of Army O-5 between officers with graduate education and those without?" By means of classification trees we also tried to uncover structure in the data set to help the reader comprehend the factors that have greatest impact on promotion.

The data used in this study was taken from the Active Duty Military Master File using fiscal years 1981 through 2001 as cohort data.

C.   SCOPE, LIMITATIONS AND ASSUMPTIONS

The data did not distinguish between fully-funded, partially-funded, and unfunded graduate education. The interesting question of whether there is a difference in promotion rates between officers having graduate degrees from different sources could not be addressed.

Some other factors that could potentially affect promotion rates, such as GPA, physical training records, awards, and second language could not be included in the study because they were not present in the data.

D. COURSE OF THE STUDY

This thesis is comprised of five chapters. Chapter II reviews pertinent literature and previous studies relevant to effects of graduate education on officer promotions. Chapter III describes the data sets and variables used for the models. It also explains the statistical models and techniques used for the study. Chapter IV consists of preliminary, multivariate and classification tree analyses. Chapter V summarizes the conclusions of the analyses and presents recommendations for further study.

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

A.   GRADUATE EDUCATION IN THE ARMY

There are two military schools specifically designated for the technical graduate education of US military officers. Naval Postgraduate School was established as a school of marine engineering at Annapolis in 1909. This small program, consisting of 10 officer students and two Navy instructors, would later become today's Naval Postgraduate School with 1,500 students coming from all service branches of the U.S. defense community and the services of more than 25 allied nations. Second technical school is the Air Force Institute of Technology, located in Dayton, Ohio. There is no particular technical graduate school for Army officers, but they can be selected to attend either of these two technical schools (Naval Postgraduate School General Catalog, 2001).

The Army War College in Carlisle, Pennsylvania and the Naval War College in Newport, Rhode Island are two other military schools primarily for strategic leadership, national security and military strategy studies. They offer a master's degree in National Security and Strategic Studies.

Other than those four military schools, officers can obtain graduate education by attending civilian universities full-time. Their curricula would not necessarily have a military focus, but the Army would sanction studies for certain programs deemed to meet its needs.

Officers who do not obtain a fully-funded or partially-funded graduate education program during which they only attend school can apply for graduate education on behalf of their own. That is called a non-funded graduate program.

Department of Defense Directive, Number 1322.10, "Policy on Graduate Education for Military Officers", describes fully-funded graduate education as when an officer receives full pay and allowances while pursuing a graduate degree, the majority of the tuition and other schooling costs being assumed or paid by the U.S. Government or by another organization. The officer attends school instead of performing usual military duties.

Partially funded graduate education is when the officer receives full payment and allowances pursuing a graduate degree, the majority of tuition and other schooling costs being paid by the officer from personal funds or benefits to which the officer was entitled. The officer attends school instead of performing usual military duties.

On the other hand, the officer pays the majority of tuition and schooling costs for unfunded graduate education. The officer attends school during off-duty time.

The difference lies in the obligatory duty after getting the master degree. As described by Army Regulation 614-100, under the title of "Policies for Assignments to Utilize Education or Experience," an officer who attends civil schooling and obtains an advanced degree under any Army Civil Schools Program or

receives fully or partially-funded support in a program of study lasting 26 weeks or more is considered an obligated Army Educational Requirements System (AERS) asset. They are required by the DOD Directive 1322.10 and AR 621-108 to serve a utilization tour in a validated AERS position for a minimum of 36 months.

B. PROMOTION IN THE ARMY

1. The Promotion Process

The Army used to maintain a single Active Duty List (ADL) on which officers were to be carried in order of seniority. They were considered for promotion each time a selection board was convened to consider officers in an established zone of consideration for their competitive category.

Since the data we used were based on the former promotion system, which grouped all officers in a single Active Duty List, the information given below is mostly related to the single active duty list procedure; however, the new promotion system under the name of "career fields" is also discussed briefly.

Within the former promotion system Title 10 USC provided a single promotion process for all officers on active duty on the ADL, regardless of their component.

Changes in authorizations, losses and promotions to the next higher grade create fluctuations in both the time in service (TIS) and time in grade (TIG) at which promotions occur. Under ideal circumstances, each qualified officer would advance through the grade structure with some degree of predictability. However, a

standardized promotion flow does not occur consistently due to expansion and contraction of the Army, changes in promotion policies and variations in officer losses each year.

The promotion timings, as stated in Department of Defense Instruction (DODI) 1320.13, are expressed in terms of years of Active Federal Commissioned Service (AFCS) at which promotion occurs.

The promotion opportunity, as stated in DODI 1320.13, is the ratio of total number of officers selected for promotion to the eligible in-the-zone population. Promotion timing and opportunity objectives are shown in the table below.

PROMOTE TO	TIMING	TIG	PRO.OPPORTUNITY
1LT/02	18 Months	18 Months	Fully Qualified
CPT/03	4 Years	2 Years	Best Qualified(90%)
MAJ/04	10 Years +/-1 Year	3 Years	Best Qualified(80%)
LTC/05	16 Years +/-1 Year	3 Years	Best Qualified(70%)
COL/06	22 Years +/-1 Year	3 Years	Best Qualified(50%)

Table 2.1.   TIS, TIG and Promotion Opportunity from Department of the Army Pamphlet 600-3.

Department of the Army (HQDA) centralized boards select officers for promotion from Captain through Colonel. Selection boards are asked to recommend best-qualified officers from an inclusive zone of consideration (ZOC). The ZOC includes officers from above, in and below

the promotion zone. When the number of officers being considered exceeds the maximum number to promote, the boards operate under best-qualified criteria. Centralized boards are provided minimum promotion requirements by branch, functional area or area of concentration to ensure the Army's skill and grade mix. Recommendations are based upon branch and functional area competency and the potential to serve in the higher. Factors considered include:

1. Performance
2. Embodiment of Army values
3. Professional attributes and ethics
4. Integrity and character
5. Assignment history and professional development
6. Military bearing and physical fitness
7. Attitude, dedication and service
8. Military and civilian education and training
9. Concern for soldiers and families.

(Department of the Army Pamphlet 600-3, "Commissioned Officer Development and Career Management," 1998)

2. Career-Field Based Management System

The Officer Personnel Management System XXI Study (1997) summarizes the reasons that a new mechanism for promotions is needed. Today's Army's increasing dependency on information-age technology requires a greater depth of knowledge and experience in functional areas. Senior officers with high-degree functional area skills become especially more important. Yet it takes time to establish essential experience and expertise in the institutional

Army functions. The fact that specialty experts can no longer afford to serve in units for long periods and that the Army can no longer afford to deny these officers promotion brings about a dilemma. Under the former system, the Army lost myriad officers with invaluable expertise and experience just because the system did not distinguish officers with regard to their background.

The Army's need, therefore, was a promotion system by which officers possessing expertise in the Army's institutional and strategic functions compete amongst themselves, rather than against officers with a war-fighting, operational background. This system brought alternative developmental and career paths for those who did not have the opportunity to serve as key leaders such as Operations Officers (S3s) or Executive Officers (XOs). These new career paths were introduced because the former promotion system grouped officers from different tracks of expertise and experience into one promotion list and thus caused invaluable officers to be lost just because they pursued non-operational tracks.

With these objectives, the Task Force established an officer management system based on career fields. Under OPMS XXI, officers will continue to have the same pattern as before until getting promoted to Major. After being promoted to Major, however, officers are grouped into management categories, career fields, with regard to their branches and functional areas. Officers will compete for promotion only against officers within the same career field, and they will be required to meet only their branch or functional area requirements. Each career field has its

own developmental track for officers, required for the readiness of the Army today and into the 21st century. A list of the career fields follows:

1.   Operations

2.   Information Operations

3.   Institutional Support

4.   Operational Support

(Officer Personnel Management System XXI Study, 1997).

3.   Lieutenant Colonel Development

Officers generally reach this rank between the 17th and 22d years of their service. Those officers selected for promotion to Lieutenant Colonel then begin their senior field grade years, where they make the maximum contribution to the Army as commanders and senior staff officers.

> Attaining the grade of Lieutenant Colonel is considered to be the hallmark of a successful career. Officers in the grade of Lieutenant Colonel serve as senior leaders and managers throughout the Army, providing wisdom, experience, vision and mentorship mastered over many years in uniform. (Department of the Army Pamphlet 600-3, "Commissioned Officer Development and Career Management," 1998)

As mentioned by the OPMS XXI Task Force, promotion to Lieutenant Colonel ensures an officer's reaching retirement eligibility, obtaining career security, and accruing the significant financial compensation and security that accompany that status. Thus the Task Force recognizes promotion to Lieutenant Colonel as a reasonable

career goal, one that is viable and credible for the majority of competent officers in the current force.[1]

The career development goals of a Lieutenant Colonel are to gain branch, functional area and skill proficiency at the senior levels through assignments and schooling. Most officers will serve in high-visibility billets in either their branch or functional area, with a possible assignment to a branch or functional area generalist position. Graduate education not only helps them to be promoted, but also, and more importantly, maintains the capabilities of the officer corps in the face of burgeoning technological advancements.

C.   PREVIOUS STUDIES

Although there have been a number of analyses on the impact of graduate education on promotion of Navy officers, there have not been many theses or studies devoted to the Army officer promotion process.

Buterbaugh (1995) proposed a multivariate model to study the effects of academic performance and graduate education on the promotion of senior US Navy officers, to the rank of Commander and Captain. Using data from the Officer Promotion History Files and categorizing the data into warfare communities and two separate time periods, he used ordinary least squares and maximum likelihood logit

---

[1] Nearly 70% of the 3,300 officers responding to the OPMS XXI Survey stated that they would like to stay in the Army beyond the minimum number of years needed for retirement (20 years). Interestingly, 36% of the respondents defined career success as achieving the rank of lieutenant Colonel, while 29% defined success as attaining the rank of Colonel.

regression models to estimate the probability of being promoted to these two ranks. The findings reveal that both undergraduate academic performance and graduate education were significantly and positively associated with the probability of promotion.

Branigan (2001) analyzed the factors associated with retention to the O-5 promotion point and selection for promotion to O-5 for Marine Corps officers. His conclusion was that the career-minded officer who chooses to participate in the Marine Corps' graduate education programs could look forward to a long, secure career and anticipate a greater chance of promotion to O-5. The results also helped alleviate perceptions that participation in graduate education programs diminishes an officer's prospects for promotion to higher ranks.

Bowman and Mehay (1998) examined the relationship between individual productivity and graduate education by analyzing the effect of graduate education on promotion to the rank of Lieutenant in the US Navy. We will paraphrase their methods and findings. They emphasized statistical correction of selectivity bias that comes with an individual's decision, and the Navy's selection of individuals, to participate in funded graduate education programs. They included variables such as college GPA, undergraduate degree and graduate education and accession source. Standard demographic characteristics such as race, sex, and marital status were also included. Single-stage estimates from their model indicated that among those reviewed for up-or-out promotion to rank O-4, promotion probabilities were 10-15 percent higher for those with any

graduate degree. For officers with degrees obtained via the Navy's full-time funded program, the differential ranged from 15 to 17 percent. However, when instruments that were uncorrelated with promotion were used to predict graduate degree status, the results suggested that a sizeable portion of the relationship between graduate education and promotion was due to unobserved attributes that lead some people both to attend graduate school and to be more promotable. The selection-corrected estimates of the promotion effect of graduate education were reduced by between 40-50%. But even after controlling for selectivity bias, officers with graduate degrees were more likely to be promoted to 0-4 than officers without graduate education.

III. DATA AND METHODOLOGY

A. VARIABLE INTRODUCTION

1. Dependent Variable

Models we will use for the study will have a dependent variable, referring to promotion, and a number of independent variables representing officers' personal characteristics and military background. The dependent variable will be a dichotomous (binary) variable (PROMOTED), which assumes a value of 1 if the officer is selected for promotion to the rank of Lieutenant Colonel (O-5) and 0 if the officer is not selected.

2. Independent Variables

Independent variables are the explanatory factors referring to officers' personal characteristics and professional backgrounds. They are grouped into the following categories: Demographics, Professional and Educational Career Traits. Demographics consist of GENDER, RACE, AGE, MARITAL STATUS, and NUMBER OF DEPENDENTS. Professional factors were EDUCATION, COMMISSION SOURCE, DOD PRIMARY OCCUPATION CODE (DPOG) and PRIOR ENLISTED (PE). Education is assigned 1 if the officer had a baccalaureate degree, and 2 if the officer had a graduate degree. Marital Status is assigned 1 if the officer is single, 2 if the officer is married, and 3 if officer is no longer married. Table 3 lists each variable and its description.

PROMOTED	0 IF OFFICER IS NOT PROMOTED
	1 IF OFFICER IS PROMOTED
GENDER	1 MALE
	2 FEMALE
RACE	0 IF UNKNOWN
	1 IF WHITE
	2 IF BLACK
	3 IF OTHER (HISPANIC, AMERICAN INDIAN, ASIAN)
MARITALSTAT	1 IF SINGLE
	2 IF MARRIED
	3 IF NO LONGER MARRIED (DIVORCED, LEGALLY SEPARATED)
NUMDEPEND	1 IF MEMBER ONLY AND 0 DEPENDENTS
	2 IF MEMBER AND 1 DEPENDENT
	3 IF MEMBER AND 2 DEPENDENTS
	4 IF MEMBER AND 3 DEPENDENTS
	5 IF MEMBER AND 4 OR MORE DEPENDENTS
EDUCATION	1 IF BACCALAURATE DEGREE OR LOWER LEVEL
	2 IF MASTER'S AND FIRST PROFESSIONALS
COMMSOURCE	0 IF UNKNOWN
	1 ACADEMY
	2 ROTC/NROTC, SCHOLARSHIP
	3 ROTC/NROTC, NONSCHOLARSHIP
	4 DIRECT APPOINTMENT, NON PROFESSIONAL
DPOG	0 IF UNKNOWN
	1 IF TACTICAL OPERATIONS OFFICER
	2 IF INTELLIGENCE OFFICER
	3 IF ENGINEERTING AND MAINTENANCE OFFICERS
	4 IF HEALTH CARE OFFICERS
	5 ADMINISTRATORS
	6 SUPPLY, PROCUREMENT AND ALLIED OFFICER
PE	N IF NOT PRIOR ENLISTED
	Y IF PRIOR ENLISTED
AGE	CONTINUOUS

Table 3.1.   Variable Names and Descriptions.

B. DATA SETS

The data used in this study were taken from the Active Duty Military Master File for fiscal years 1981 through 2001 as cohort data sets. Data sets originally had 460 columns (constant variables over time like sex and enlisted status or time dependent variables like age and education) of longitudinal information. Besides yearly

tracked ones, from each officer's active duty year through 2001, there were variables pertaining to unit management, military occupation, career timeline events, and military and personal demographics.

Variables related to officers' personal demographics, military occupational information and educational levels were obtained from the original data sets. First, for each cohort, warrant officers and officers having rank other than O-1 at the beginning year of active duty were eliminated. Second, to prevent any bias from early resignations, only officers who reached the rank of O-3 were selected for the analysis.

With respect to the variables tracked yearly, such as Marital Status or Educational Level, status at the tenth-year point was selected to represent each of the officers. There were two reasons for choosing the tenth-year point as the reference for yearly tracked information. The fact that the average promotion time to the rank of Major (O-4) is ten years was the first reason. The second reason was that almost all cohorts had a considerable number of officers with graduate education by that time. To produce the response variable PROMOTED, pay grade information was used to distinguish officers who were promoted to rank O-5. Since being promoted to rank O-5 requires an average of 17 years, only the first three cohorts were used. A pooled data set, the combination of three separate cohort data sets, was also used.

Once all necessary removals were made, each column of information was examined and necessary factorizing and decoding were established. For instance, the DOD Primary

Occupation Code was coded numerically in the original data sets, but for the purpose of the thesis research they were decoded; their references were searched from the DOD Occupational Data Base, and the codes were then categorized for the modeling process.

Only the first three cohort data sets, (1981, 1982, and 1983) were organized for the study. Cohort 1981 had 2653 observations; cohort 1982 had 2274 observations; and cohort 1983 had 1907 observations. All data sets consisted of 10 variables, 9 of which were categorical. Promotion rates were 40%, 43.2%, and 52.5% for the three cohorts respectively. Officers in each of the cohorts were predominantly white, male and married. Rates of officers with graduate degree were 27.7%, 28.4%, and 29.9% respectively. Rates of Prior Enlisted were 20.4%, 29.7%, and 28.1% respectively. The average age was approximately 34 for all data sets. The pooled data set was constructed by combining all three data sets into one. This data set had a 44.8% promotion rate.

C.   METHODOLOGY

1.   Logit Regression

The goal of an analysis using this method is the same as that of any model-building technique used in statistics: to find the best-fitting and most parsimonious and reasonable model by which to describe the relationship between an outcome (dependent or response) variable and a set of independent (predictor or explanatory) variables. What distinguishes a logistic regression model from linear

regression is that the outcome variable in logistic regression is binary or dichotomous.

In any regression model the key quantity is the mean value of the outcome variable, given the value of the independent variables. This quantity is called the conditional mean and is expressed as " $E(Y|X)$ " where Y denotes the outcome variable and X denotes the value of the independent variables. In linear regression we assume that this mean may be expressed as an equation linear in x, such as " $E(Y|X) = \beta_0 + \beta_1 X$ ". The specific form of the logistic regression model we used, which formulates the conditional mean of the regression equation to be bounded between zero and one, is as follows:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}} \ .$$

($\pi(x)$ Refers to $E(Y|X)$ to simplify the notation)

The Logit function, that is the natural logarithm of the odds ratio, is defined in terms of $\pi(X)$ as follows:

$$g(X) = \ln\left[\frac{\pi(X)}{1 - \pi(X)}\right] = \beta_0 + \beta_1 X \ \ .$$

The importance of this transformation is that $g(X)$ has many of the desirable properties of a linear regression model. The logit, $g(X)$, is linear in its parameters, may be continuous, and may range from negative infinity to positive infinity.

The second important difference between the linear and logistic regression models concerns the conditional distribution of the outcome variable. In the linear

regression model we assume that an observation of the outcome variable may be expressed as $y = E(Y|x) + \varepsilon$. The most common assumption is that ε follows a normal distribution with mean zero and some constant variance. This is not the case with a dichotomous outcome variable. We express the value of the outcome variable given x as $y = E(Y|x) + \varepsilon$. Here the quantity ε may assume one of two possible values. If $y = 1$ then $\varepsilon = 1 - \pi(x)$ with probability $\pi(x)$, and if $y = 0$ then $\varepsilon = -\pi(x)$ with probability $1 - \pi(x)$. Thus the outcome variable follows a binomial distribution with probability $\pi(x)$ given by the conditional mean (Hosmer and Lemeshow, 1989, p.7).

In our study, $g(x)$ will be a linear function of the vector of independent variables, x_i. Then the probability $\pi(x)$ is a nonlinear function with an S-shape. Estimates of β values, denoted by b, are hard to interpret because of the transformation. They represent the slope or rate of change of the logit of the dependent variable per unit of change in the independent variable. Interpretation involves two issues: determining the functional relationship between the dependent variable and the independent variable, and appropriately defining the unit of change for the independent variable. The estimated probability that the response variable (promotion in our study) takes on can be obtained from the formula $\hat{\pi}(x) = 1/(1 + \exp(-\sum x_i b_i))$ where b_i is the estimated coefficient value obtained for the ith variable from the logistic model. Other variables being constant, a change of one

unit in the variable's value multiplies the odds ratio (explained in Chapter IV) by $\exp(b_i)$.

2. Classification Tree Modeling

Tree-based models provide an alternative to linear and logistic models for classification problems. The models are fit by binary recursive partitioning whereby a dataset is successively split into increasingly homogeneous subsets until it is infeasible to continue. Tree-based modeling is an exploratory technique for uncovering structure in data, increasingly used for assessing the adequacy of linear models and summarizing large multivariate datasets. The rules are determined by a procedure known as recursive partitioning (Breiman *et al.*, 1984).

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ANALYSIS

A.   PRELIMINARY DATA ANALYSIS

Table 4.2 provides descriptive statistics for all data sets as percentages except for age. Of the 7583 officers from the pooled data set who reached at least rank O-3, only 3400 (44.8%) were promoted to O-5; 28.94% had graduate education from any source, and 27% had been enlisted before. The number of observations for the 1981, 1982, and 1983 data sets were 3091, 2529, and 2260 respectively. The proportions of officers with graduate education were 27.7%, 28.4%, and 29.9% respectively.

Table 4.1 shows the results of Chi-Square (for factorial variables) and t-Test (for continuous variables) statistics, which were used to compare each variable across two groups: officers who were selected for promotion and officers who were not selected. The p-values for all variables except for DPOG were smaller than 0.05 significance level. The results are evidence that all variables except for DPOG differ between the two officer groups.

VARIABLE	CHI-SQUARE and T-TEST P-VALUES
EDUCATION	0
GENDER	0.0455
RACE	0.0084
MARITALSTAT	0.0409
NUMDEPEND	0.0027
COMMSOURCE	0
DPOG	0.6819
PE	0
TENTH.YEAR.AGE	0

Table 4.1.   P-Values for Variables Across Two Groups of Officers (Promoted and Not Promoted to O-5).

Cohort	Factor							
1981 COHORT	**PROMO**	0:60	1:40					
	GENDER	1:88.5	2:11.5					
	RACE	1:90.4	2: 5.4	3: 4.2				
	MAR.STAT	1:11.1	2:84.4	3: 4.4				
	NUMDEP	1:16.5	2:18.4	3:20.4	4:29.8	5:14.9		
	EDUCATION	1:72.3	2:27.7					
	COMSOURCE	0:41.7	1:19.8	2:18.7	3:13.3	4: 6.7		
	DPOG	0:41.1	1:30.9	2: 3.7	3:15.3	4: 1.7	5: 6.8	6: 0.5
	PE	N:79.6	Y:20.4					
	AGE	33.84						
1982 COHORT	**PROMO**	0:56.8	1:43.2					
	GENDER	1:88.4	2:11.6					
	RACE	1:68.9	2: 8.4	3: 2.8				
	MAR.STAT	1:11.1	2:83.8	3: 5.1				
	NUMDEP	1:16.9	2:18.9	3:19.3	4: 29	5:15.9		
	EDUCATION	1:71.6	2:28.4					
	COMSOURCE	0: 52	1:17.9	3:13.3	4: 3.1			
	DPOG	0:37.6	1:30.8	2: 4	3:14.5	4: 2.5	5: 8	6: 2.6
	PE	N:70.3	Y:29.7					
	AGE	34.15						
1983 COHORT	**PROMO**	0:47.5	1:52.5					
	GENDER	1:89.4	2:10.6					
	RACE	0: 37	1:53.6	2: 5.8	3: 3.7			
	MAR.STAT	1: 9.9	2:86.2	3: 3.9				
	NUMDEP	1:15.7	2:18.8	3:21.2	4:29.3	5:14.9		
	EDUCATION	1:70.1	2:29.9					
	COMSOURCE	0:37.7	1:16.9	2:20.9	3:19.7	4: 4.8		
	DPOG	0:38.9	1:32.6	2: 5.3	3:10.4	4: 4.2	5: 6.3	6: 2.3
	PE	N:71.9	Y:28.1					
	AGE	34.16						
POOLED DATA	**PROMO**	0:55.2	1:44.8					
	GENDER	1:88.7	2:11.3					
	RACE	0:15.4	1: 74	2: 6.9	3: 3.7			
	MAR.STAT	1:10.7	2:84.8	3: 4.5				
	NUMDEP	1:16.4	2:18.9	3:20.3	4:29.4	5:15.1		
	EDUCATION	1:71.1	2:28.9					
	COMSOURCE	0:42.1	1:20.6	2:16.6	3:14.8	4: 5.9		
	DPOG	0:38.6	1:31.7	2: 4.4	3:13.4	4: 3.2	5: 6.9	6: 1.7
	PE	N:73	Y:27					
	AGE	34.13						

Table 4.2.   Descriptive Statistics for Each Data Set
(Values are percentages of each factor level)

Descriptive statistics were examined for each year to disclose any relationship between education level and promotion rates. Figure 4.1 indicates that the officers with graduate degrees have higher promotion probabilities throughout the years in question. Similarly Figure 4.2 indicates that officers who graduated from the US Military Academy have higher promotion probabilities than Non-Academy graduate officers. Provided that we had enough detailed information, an analysis could have been done of officers who obtained graduate degrees from different sources to examine if those sources have an effect on promotion rates.
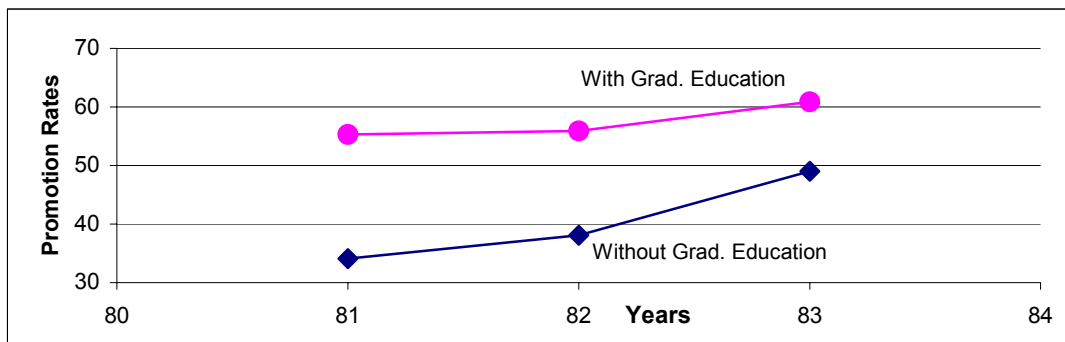

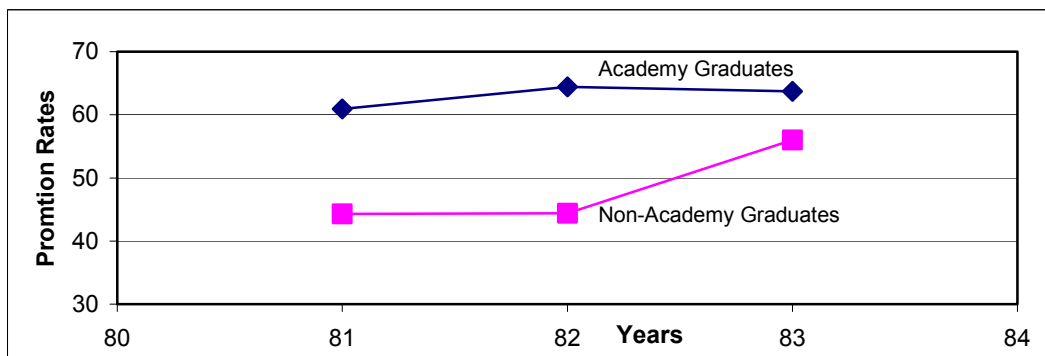Figure 4.1.   Promotion Rates for Officers With and Without Graduate Degree.

Figure 4.2.   Promotion Rates for Academy and Non-Academy Graduate Officers.

27

B.   MULTIVARIATE ANALYSIS

Multivariate modeling analyzes the effects of individual independent variables on the response variable by holding the effects of other variables constant. This thesis used the logit link function for the promotion models with the binary response variable. Initial models were used to identify those predictor variables that were significant with respect to a 95% confidence level. A stepwise model selection procedure was used to determine if any two-way interactions were significant before eliminating any main effects.

The software package S-Plus® 2000 (MathSoft, 2000) was used to estimate regression models and classification trees. After performing stepwise addition and deletion of terms, an analysis of deviance (Mc Cullagh and Nelder, 1989) test was used to determine whether the main factors or interactions are statistically significant. Having developed the models, diagnostics were checked and necessary changes were made to achieve in the final model a balance of simplicity and fit. A graph of Cook's distance versus predicted probability was used to find the most influential observations on the model. A graph of Pearson chi-square versus predicted probability was utilized to find the poorly fitted observations in the models.

Having developed the models, predictions were calculated. Both original and different data sets were used for prediction to eliminate the learning effect of evaluating the model using the same data on which they were built. The threshold for the promotion predictions

was 0.5, that is, any officer with predicted promotion probability greater than 0.5 was predicted to be promoted. The accuracy of predictions was lower than expected. The cross-validation results and model accuracy percentages are tabulated below.

Model	Prediction Accuracy	Cross-Validation Results
1981 cohort	64.5%	65.7%
1982 cohort	66.5%	65.4%
1983 cohort	63.8%	62.4%
Pooled data set	65.3%	64.5%

Table 4.3   Model Prediction Accuracy Results.

Although the misclassification rate seemed quite high, perhaps due to the lack of enough professional and educational background about officers, the results were deemed reliable. Besides inadequate information in the data sets, developments after the tenth-year point, the reference point for yearly tracked information, could certainly have an effect on promotion probabilities. The reason for choosing the tenth year point as a reference for the yearly tracked variable was to have enough observations of officers with graduate education.

Although the accuracy of individual promotion predictions was low, the reliability of models could be examined by comparison of predicted and actual group promotions. Observations were sorted and grouped with respect to their predicted probabilities; observations with predicted promotion probability between 0 and 0.1 fell into the first group, observations with prediction of

promotion probability between 0.1 and 0.2 fell in the second, and so on. Each group's average predicted promotion probability was compared to the actual promotion proportion of the same observation group. The results, as shown in Table 4.4 are close to each other within 3%, except for the first group. Figure 4.3 shows an almost diagonal line, indicating that the model can accurately predict promotions within a group.

GROUP	ACTUAL AVERAGE PROMOTION PRO.	PREDICTED AVERAGE PROMOTION PRO.
1	0.000	0.077
2	0.147	0.150
3	0.252	0.258
4	0.377	0.351
5	0.471	0.452
6	0.578	0.558
7	0.616	0.655
8	0.705	0.745

Table 4.4.   Comparison of Actual and Predicted Group
             Promotion Probabilities.
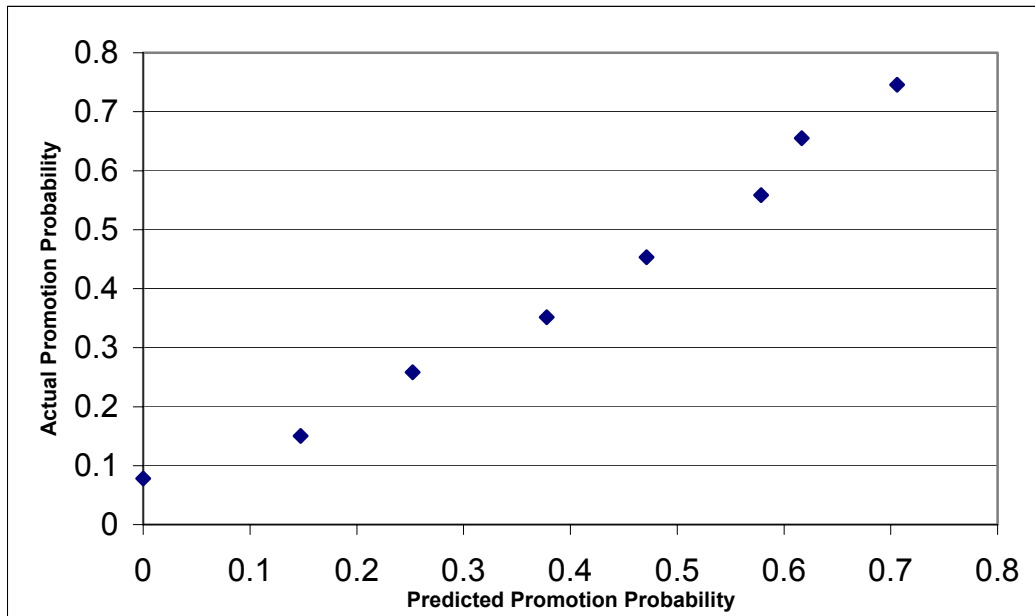

Figure 4.3.   Comparison of Actual and Predicted Group
              Promotion Probabilities.

To clarify that the models fit the data well, the Hosmer-Lemeshow test was utilized. This test uses a

30

goodness-of-fit statistic, $\hat{C}$, which is obtained by calculating the Pearson chi-square statistic depending on the observed and predicted frequencies. They demonstrated that when the number of unique observations is equal to the total number of observations and the logistic regression is the correct model the distribution of the statistic $\hat{C}$ is approximately chi-square (Hosmer and Lemeshow, 1989, p.140).

The requirements for usage of Hosmer-Lemeshow test are met by our data sets. The Hosmer-Lemeshow test produced high p-values for each model indicating that we cannot reject the null hypothesis that the models fit well. This conclusion buttressed our previous comparison of group promotions between actual and predicted values.

Table 4.5 presents the results of the model constructed for the 1981 cohort data set. Other models are included in Appendix A. Tables include the coefficient value, standard error and t-value for each variable and a confidence interval for the odds ratio, explained later on. These tables include the variables that are statistically significant at the 95% confidence level. Variables not appearing in the table are the ones not significant at 95% confidence level.

The 1981 Cohort model proposed that log-odds of promotion depend on the following terms:

RACE, MARITALSTAT, EDUCATION, COMMSOURCE, PE, TENTH.YEAR.AGE and the interaction of MARITALSTAT and PE.

	Value	Std. Error	t value	95% CI of Odds-Ratio
(Intercept)	2.262	0.880	2.571	
RACE2	-0.322	0.200	-1.609	0.489 - 1.072
RACE3	-0.456	0.232	-1.965	**0.401 - 0.998**
MARITALSTAT2	0.603	0.154	3.897	**1.349 - 2.477**
MARITALSTAT3	0.413	0.280	1.476	0.873 - 2.617
EDUCATION	0.809	0.095	8.520	**1.865 - 2.707**
COMMSOURCE1	1.189	0.125	9.485	**2.570 - 4.203**
COMMSOURCE2	0.808	0.124	6.495	**1.758 - 2.864**
COMMSOURCE3	0.712	0.134	5.302	**1.566 - 2.652**
COMMSOURCE4	0.352	0.181	1.944	**0.997 - 2.028**
PE	0.735	0.345	2.127	**1.059 - 4.107**
TENTH.YEAR.AGE	-0.117	0.025	-4.525	**0.845 - 0.935**
MARITALSTAT2PE	-0.913	0.359	-2.537	0.439 - 1.800
MARITALSTAT3PE	-0.203	0.564	-0.360	0.269 - 2.468

Table 4.5.   Logit Regression Model Summary for
              1981 Data Set.

The features of the baseline officer which S-Plus uses as the base level to compare the different factor levels of each variable were:

GENDER   male

RACE   unknown

MARITAL STATUS   single

NUMBER OF DEPENDENTS   member only

EDUCATION   no graduate education

COMMISSION SOURCE   unknown

DOD PRIMARY OCCUPATION GROUP   unknown

PRIOR ENLISTED   no

According to the 1981 cohort model, officers with a graduate degree, who were white, married, graduated from Academy rather than ROTC program and prior enlisted are most likely to be promoted. When it comes to the effect

of age, the model showed that promotion probability decreases with an increase in age.

As a quantitative example, for instance, the odds ratio associated with a graduate degree is exp(0.809)= **2.25.** This value, being greater than one, indicates a positive impact of Graduate Education on promotion.

In each model, the Education variable was statistically significant even at alpha=0.01 level.

For example, according to the 1981 cohort model, a white, married, Academy graduate, not prior enlisted, age 33, with a graduate degree has log odds of promotion given by

2.26 + 0 + 0.60 + 0.81 + 1.19 + 0 + 33*(-0.12) = 0.90

The predicted promotion probability of this officer was $[1 + \exp(-0.90)]^{-1}$ = **0.71.**

Another officer having the same features except without a graduate degree would have a different probability of promotion, delineated below:

2.26 + 0 + 0.60 + 0 + 1.19 + 0 + 33*(-0.12) = 0.18

$[1 + \exp(-0.18)]^{-1}$ = **0.55.**

According to the model, the presence of graduate education increases the probability of promotion for the officer whose traits are given above by 16%. This increase in promotion probability is not constant for the whole population, because the effect of education depends on the other variables as well.

However, by looking at the odds ratios we can argue that graduate education has a noticeable positive impact

on officer promotions to O-5. Variables whose odds ratio intervals do not include one have significant impact on the response variable.

Figure 4.4 illustrates the effects of graduate education and age on promotion probabilities. Each line shows the predicted probability of promotion versus age, one line showing officers with graduate education, and the other showing officers without graduate education. Variable Age had the same range as in the original data set. Other variables were assigned as the most common levels of each variable: male, white, married, three dependents, academy graduate, tactical operations officer, and not prior enlisted.

Figure 4.4 indicates that, under the model, the probability of promotion declines with an increase in age. Moreover, the difference between the two lines indicates a noticeable positive effect of graduate education on promotion probabilities.
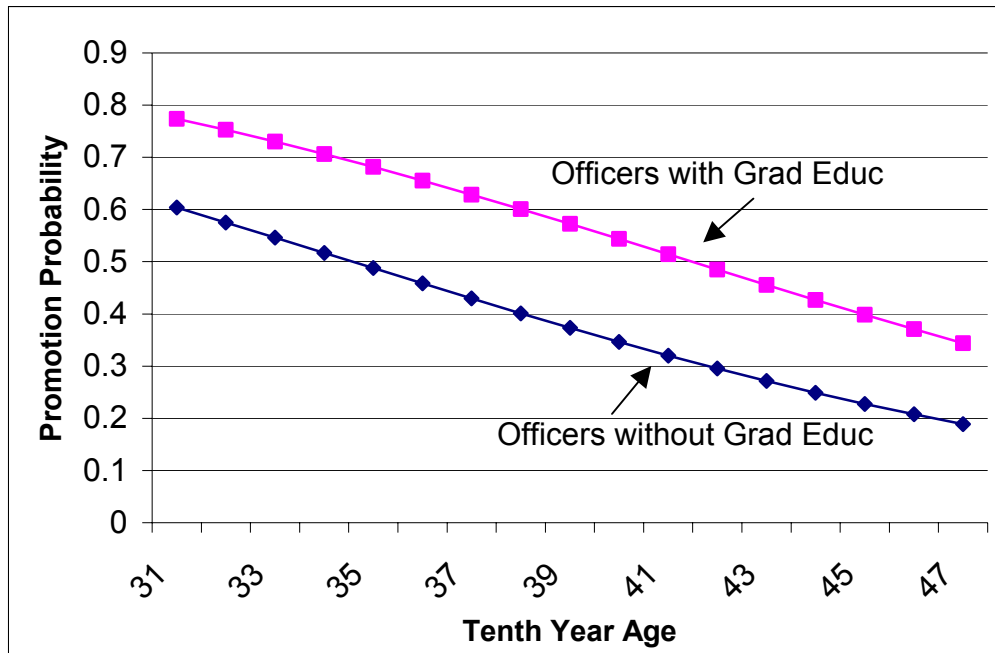
Figure 4.4.   Effect of Graduate Education and Age on Promotion Probability.

Age has a decremental effect on promotion probability. The results of the model suggest that each additional year of age subtracts 0.12 from the log odds of promotion, or multiplies the odds by exp(-0.12)=0.89. In other terms with each additional year of age, the odds of promotion decline by 11%.

Confidence intervals for the odds of each variable are also calculated and included in the tables for each model. The odds ratio confidence intervals for Education for each model are shown below.

Year	95% Confidence Interval
1981	(1.87 - 2.70)
1982	(1.73 - 2.60)

1983   (1.29 - 1.99)

Pooled data   (1.79 - 2.25)

Thus, having a graduate education increases the odds of promotion by a factor between 1.79 and 2.25 in the pooled data model. Moreover no interval includes one, that is, in each model Education has a significant positive effect on the response variable.

We interpret this result as follows:

let a = P(promoted | with grad.education)

and b = P(not promoted | with grad.education)

and A = a/b = Odds(with grad.education);

let c = P(promoted | without grad.education)

and d = P(not promoted | without grad.education)

and B = c/d = Odds(without grad.education).

Then the odds ratio for promotion is: **C = A/B.**

The odds ratio confidence intervals (CI) included in the tables are the CI of the value C calculated above. Although the estimate of the odds ratio has a skewed distribution, for large enough sample sizes it will have normal distribution. The odds ratio confidence intervals are obtained by first calculating the endpoints of confidence intervals for the coefficients and then exponentiating them as in the expression $\exp[b_i \pm Z_{1-\alpha/2} * \hat{SE}(b_i)]$ (Hosmer and Lemeshow, 1989, p.44).

As far as the other variables are concerned, being African American decreases the odds of promotion by a factor of between approximately 0.51 and 0. The corresponding range for other race, neither White nor Black, is 0.60 and 0.01.

With respect to Marital Status, being married increases the odds of promotion by a factor of 1.35 to 2.48.

Although the Commission Source variable has the Unknown source as the baseline, the differences of the odds between Commission Source 1, Academy graduate, and the other Commission Sources are conspicuous. Academy graduates have a greater odds of promotion, almost twice as high in odds ratio, compared to other sources (ROTC and Direct Appointment).

Serving as an enlisted soldier before being promoted to the Officer ranks increases the odds of promotion by a factor of between 1 and 4.1, according to the 1981 cohort model.

C.   CLASSIFICATION TREE ANALYSIS

A classification tree analysis comprises a set of model-free methods for analyzing multivariate data (Fisher and Lenz, 1996; Biggs et al., 1991; Cox, 1989) and mining large databases for useful knowledge (Elder and Pregibon, 1996).

Classification trees have nodes representing questions and have branches (arcs) at each node representing possible answers. Internal nodes are also

37

called splits, while leaf nodes (tips of the tree) represent probabilistic classifications or predictions of the value of the dependent variable. The tree stops growing when no additional questions improve the ability to predict the value of the dependent variable, as measured by the selected criterion. The best split is the one that produces the largest decrease in diversity of the classification label within each partition. In other words, the algorithm tries to increase homogeneity. Trees may be grown larger than needed and then pruned back until the estimated error rate is minimized (Breiman *et al.*, 1984).

Classification tree modeling was used to explore the relationship between the independent factors and promotion probability. The tree graph presents all information in a simple, straightforward way, and allows one to digest the information in much less time. Terminal nodes are the most homogenous groups of observations and they give the predicted values for the dependent variable, PROMOTED.

Since S-Plus builds trees that over-fit the original data set by growing larger trees than needed, the optimal size of the tree was found with respect to the deviances by means of cross-validation. Cross-validation splits the data set into ten different groups, and tree models are grown by leaving each of the subsets out in turn. The subset not used for tree building is used for prediction with the tree. All different tree models' deviances are calculated within a range of tree sizes. The tree sizes and corresponding deviances are graphed and S-Plus gives

the tree size that has the minimum deviance. Trees were pruned to the best size, based on cross-validation.

Figure 4.5 shows the classification tree for the pooled data set. The predicted value for the response variable, the probability of promotion, is centered in each node. The number under each node is the misclassification error rate; the numerator gives the number of wrongly predicted observations, and the denominator gives the total number of observations in that node. Rectangular nodes are terminal nodes, used for prediction.

Besides being a tool by which to unfold the structure of the data set, the tree model can be used for individual predictions as well. The first split separates the officers with regard to their Commission Sources. This split suggests that Academy and ROTC/SCHOLARSHIP graduates have higher chance of promotion than officers who are Direct Appointed and ROTC/NONSCHOLARSHIP graduates. The difference in promotion probabilities between these two groups is noticeable, 23%.

The second split indicates that officers with graduate education within Commission Sources 1 and 2 have a higher chance of being promoted than officers without a graduate degree. The difference in promotion probability between these two groups is a little bit higher than 16% in favor of those with graduate degree.
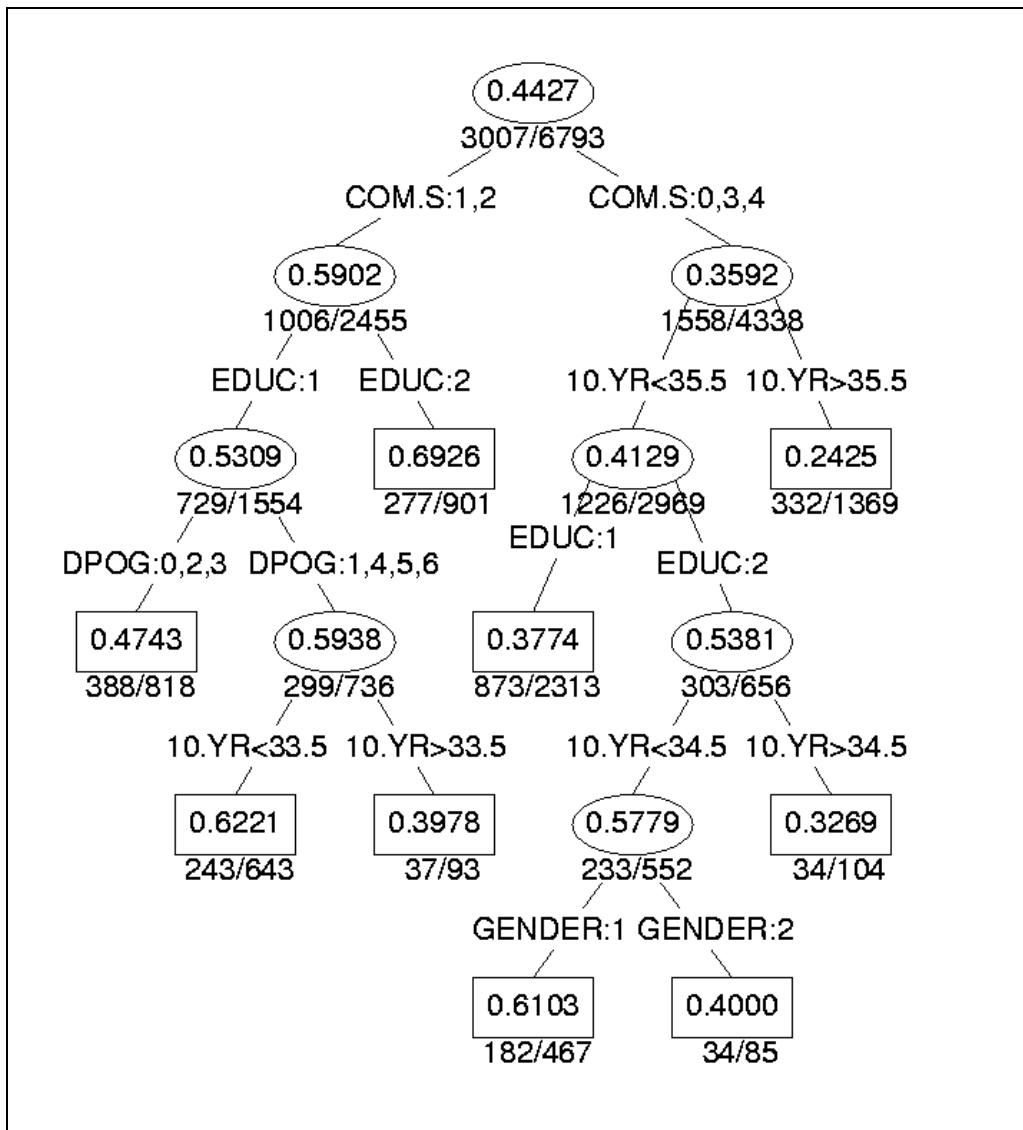
Figure 4.5.   Classification Tree for Pooled Data Set.

Among officers with Commission Source 1 or 2, Education 1, and DPOG 1, 4, 5, or 6, age makes a difference in promotion probability. Officers in this group younger than 33.5 have a 22% higher promotion probability than officers older than 33.5.

Classification trees estimated from the other data sets also indicated that Education level, Commission

Source and Age are the primary predictors of promotion. The positive effect of graduate education on promotions, under this tree model, ranges between 16% and 20% depending on the other predictors. For Commission Source the same effect is between 11% and 14%.

Misclassification rates and individual predictions for the tree models were consistent with those of logistic regression.

The prediction accuracies for the tree models are:

1981 cohort year data set: 63.7%

1982 cohort year data set: 64.9%

1983 cohort year data set: 62.8%

Pooled data set   : 64.5%.

THIS PAGE INTENTIONALLY LEFT BLANK

V. SUMMARY, LIMITATIONS AND RECOMMENDATIONS

The purpose of this study was to examine if having a graduate degree benefits Army officers in being promoted to rank O-5. This purpose could also have included distinguishing any differences between officers who acquired their graduate degrees from different sources and by different means: civilian or military schools and fully, partially or non-funded programs. But because the data did not contain this information, our focus was to analyze the effect of graduate education on promotions as a whole, and scrutinize other factors' impacts on the outcome variable as well.

In order to have a robust model with which to predict or extrapolate, it is vital to have enough predictor variables related to the response variable. Some data on professional and educational background were lacking; models were built with the data at hand. Even having ample predictor variables would not suffice to estimate individual behavior with high accuracy. Misclassification rates found by our models, affected both by the uncertainty of behavioral factors and by the lack of data, were lower than expected. However, predictions of trends for groups of people did give us a general appreciation of model performance for people with similar traits.

In light of the coefficients produced by the logit models, the short answer to the main question of the thesis is "Yes, graduate education is associated with higher probability of promotion to the rank of Army O-5." In quantitative terms the odds ratio for graduate

education lies between 1.79 and 2.25. This ratio suggests that promotion ratio among officers with graduate degrees is 1.79-2.25 times the same ratio for officers without a graduate degree. We cannot assert a general difference between these two groups of officers in terms of probability of promotion. However, as mentioned in the previous chapter, graduate education grants a higher promotion probability for officers who have similar traits.

Classification tree models confirm the positive impact of graduate degree on officer promotions. Commission Sources also seemed to be related to differences in promotion predictions. Officers graduated from Academy or ROTC / Scholarship have higher promotion probabilities compared to those from other sources.

Besides having graduate education, being an Academy or ROTC/NROTC SCHOLARSHIP graduate or being married increases the odds ratio for officer promotion. On the other hand, age has a negative effect on promotion, that is, the promotion probability decreases with an increase in age. Number of dependents, DOD primary occupation code, gender, and prior enlisted variables seemed not to be statistically important.

A more detailed analysis by comparing the different source and different type of graduate education could be done if data were available. Also adding the fitness reports, undergraduate majors, and GPA to the data sets may increase the accuracy of the models.

A follow-on study can be done in the future to compare the promotion probabilities between former

military branches (DPOG) and new career fields. The main purpose of this kind of study is to examine if the new system solves the problem of denying specialty experts promotion as in the former system. Furthermore, this study can also scrutinize the impact of graduate education separately on promotion of officers within operational branches and within non-operational specialty fields. It would be essential to gather as much information as possible about officers' career fields, undergraduate majors and GPA's, graduate education status (type and source of graduate education), commissioning sources, fitness reports, and demographic traits.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. LOGISTIC REGRESSION MODEL RESULTS

The 1982 Cohort model proposed that log-odds of promotion depend on the following terms:

RACE, MARITALSTAT, NUMDEPEND, EDUCATION, COMMSOURCE, PE, TENTH.YEAR.AGE, and the interaction of RACE and NUMDEPEND.

Coefficients	Value	Std. Error	t value	95% CI of Odds-Ratio
(Intercept)	5.938	0.909	6.534	
RACE1	-0.928	0.302	-3.073	0.218 - 0.714
RACE2	-1.301	0.462	-2.814	0.109 - 0.673
RACE3	-1.126	0.714	-1.578	0.080 - 1.313
MARITALSTAT2	0.868	0.221	3.920	1.543 - 3.676
MARITALSTAT3	0.266	0.278	0.958	0.757 - 2.247
NUMDEPEND2	-0.979	0.373	-2.623	0.180 - 0.780
NUMDEPEND3	-1.838	0.392	-4.684	0.073 - 0.343
NUMDEPEND4	-0.953	0.355	-2.684	0.192 - 0.773
NUMDEPEND5	-1.000	0.387	-2.588	0.172 - 0.784
EDUCATION	0.753	0.103	7.288	1.733 - 2.599
COMMSOURCE1	0.723	0.139	5.198	1.733 - 2.599
COMMSOURCE2	0.362	0.140	2.594	1.092 - 1.889
COMMSOURCE3	0.185	0.143	1.298	0.910 - 1.590
COMMSOURCE4	0.308	0.285	1.082	0.778 - 2.378
PE	-0.245	0.130	-1.881	0.606 - 1.010
TENTH.YEAR.AGE	-0.186	0.026	-7.060	0.788 - 0.874
RACE1NUMDEPEND2	0.728	0.393	1.854	
RACE2NUMDEPEND2	0.645	0.646	0.998	
RACE3NUMDEPEND2	0.790	0.960	0.823	
RACE1NUMDEPEND3	1.551	0.407	3.811	
RACE2NUMDEPEND3	2.474	0.618	3.999	
RACE3NUMDEPEND3	-2.623	3.986	-0.658	
RACE1NUMDEPEND4	1.048	0.363	2.888	
RACE2NUMDEPEND4	1.640	0.591	2.776	
RACE3NUMDEPEND4	1.422	0.881	1.614	
RACE1NUMDEPEND5	0.970	0.406	2.389	
RACE2NUMDEPEND5	-0.073	0.765	-0.095	
RACE3NUMDEPEND5	1.364	0.961	1.419	

Table A.1.   Logit Regression Results for 1982 Data Set.

The 1983 Cohort model proposed that log-odds of promotion depend on the following terms:

MARITALSTAT, NUMDEPEND, EDUCATION, COMMSOURCE, PE, TENTH.YEAR.AGE, and the interactions of COMMSOURCE and PE, PE and TENTH.YEAR.AGE.

Coefficients	Value	Std. Error	t value	95% CI of Odds-Ratio
(Intercept)	5.948	0.851	6.990	
MARITALSTAT2	1.023	0.247	4.140	1.714 – 4.517
MARITALSTAT3	-0.027	0.327	-0.084	0.511 – 1.849
NUMDEPEND2	-0.409	0.226	-1.805	0.425 – 1.035
NUMDEPEND3	-0.799	0.226	-3.524	0.288 – 0.701
NUMDEPEND4	-0.460	0.223	-2.057	0.406 – 0.978
NUMDEPEND5	-0.393	0.242	-1.620	0.419 – 1.085
EDUCATION	0.470	0.110	4.254	1.288 – 1.987
COMMSOURCE1	0.221	0.158	1.394	0.914 – 1.704
COMMSOURCE2	0.353	0.152	2.325	1.057 – 1.918
COMMSOURCE3	-0.054	0.174	-0.310	0.673 – 1.332
COMMSOURCE4	0.670	0.350	1.909	0.982 – 3.888
PE	-0.518	0.177	-2.916	0.420 – 0.843
TENTH.YEAR.AGE	-0.190	0.025	-7.600	0.786 – 0.868
COMMSOURCE1PE	0.069	0.866	0.080	0.196 – 5.863
COMMSOURCE2PE	0.258	0.399	0.647	0.592 – 2.832
COMMSOURCE3PE	0.936	0.275	3.401	1.487 – 4.374
COMMSOURCE4PE	-0.170	0.499	-0.342	0.317 – 2.242

Table A.2.   Logit Regression Results for 1983 Data Set.

The Pooled data set model proposed that log-odds of promotion depend on the following terms:

RACE, MARITALSTAT, EDUCATION, COMMSOURCE, TENTH.YEAR.AGE.

Coefficients	Value	Std. Error	t value	95% CI of Odds-Ratio
(Intercept)	4.518	0.430	10.489	
RACE1	-0.364	0.071	-5.109	0.604 - 0.798
RACE2	-0.475	0.122	-3.891	0.489 - 0.789
RACE3	-0.563	0.156	-3.452	0.429 - 0.792
MARITALSTAT2	0.525	0.085	6.128	1.429 - 1.999
MARITALSTAT3	0.099	0.151	0.656	0.820 - 1.188
EDUCATION	0.698	0.057	12.055	1.795 - 2.252
COMMSOURCE1	0.877	0.078	11.193	2.062 - 2.805
COMMSOURCE2	0.573	0.074	7.695	1.533 - 2.053
COMMSOURCE3	0.431	0.0764	5.642	1.325 - 1.789
COMMSOURCE4	0.333	0.126	2.627	1.088 - 1.789
TENTH.YEAR.AGE	-0.160	0.012	-12.752	0.830 - 0.873

Table A.3.   Logit Regression Results for Pooled Data Set.

THIS PAGE INTENTIONALLY LEFT BLANK

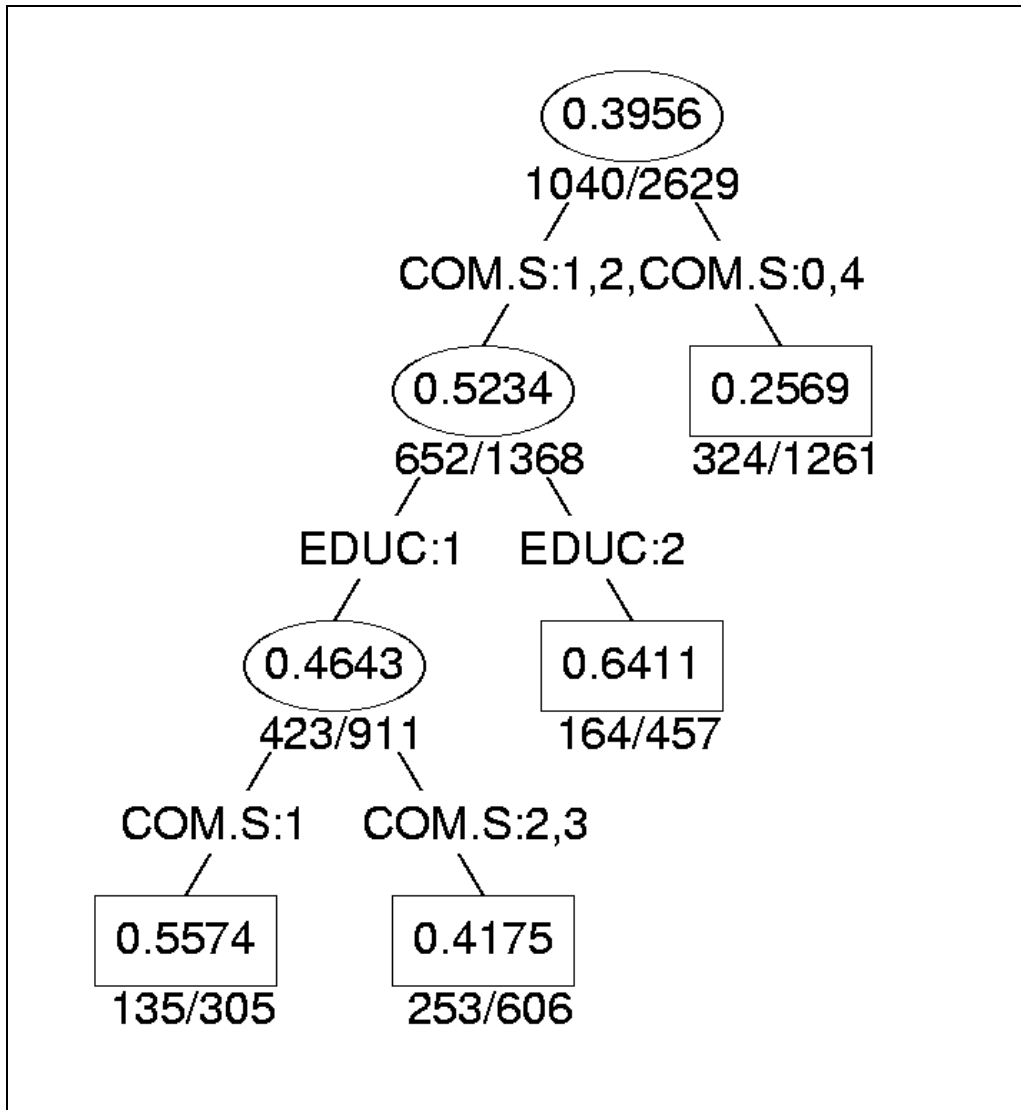# APPENDIX B. CLASSIFICATION TREE MODELS

Figure B.1.   Classification Tree for 1981 Cohort Data
            Set.

Variables actually used in tree construction:

COMMSOURCE, EDUCATION

Number of terminal nodes: 4

Residual mean deviance: 1.248 = 3276 / 2625
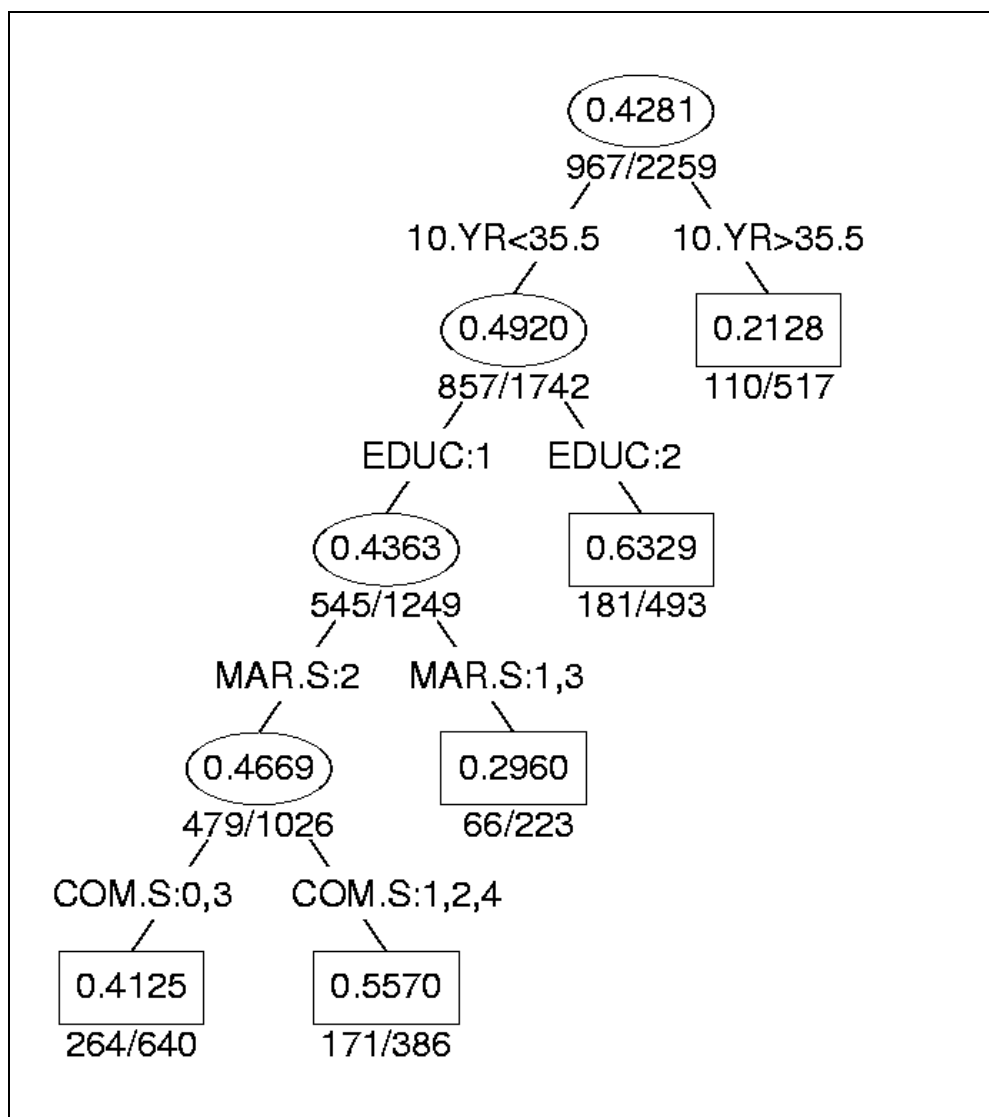
Misclassification error rate: 0.3332 = 876 / 2629

Figure B.2.   Classification Tree for 1982 Cohort Data
               Set.

Variables actually used in tree construction:

TENTH.YEAR.AGE, EDUCATION, MARITALSTAT, COMMSOURCE

Number of terminal nodes: 5

Residual mean deviance: 1.265 = 2852 / 2254

Misclassification error rate: 0.3506 = 792 / 2259
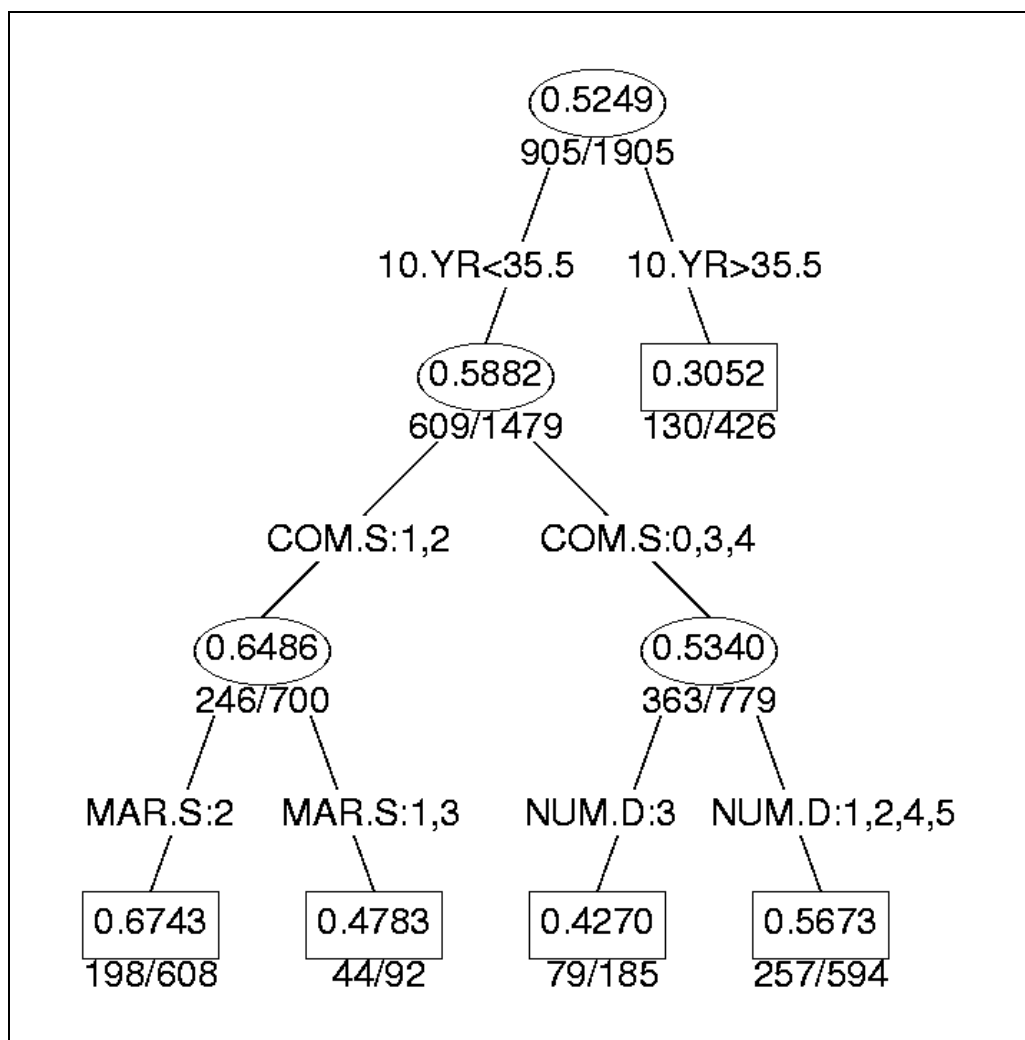
Figure B.3.   Classification Tree for 1983 Cohort Data
               Set.

Variables actually used in tree construction:

TENTH.YEAR.AGE, COMMSOURCE, MARITALSTAT, NUMDEPEND

Number of terminal nodes: 5

Residual mean deviance: 1.307 = 2484 / 1900

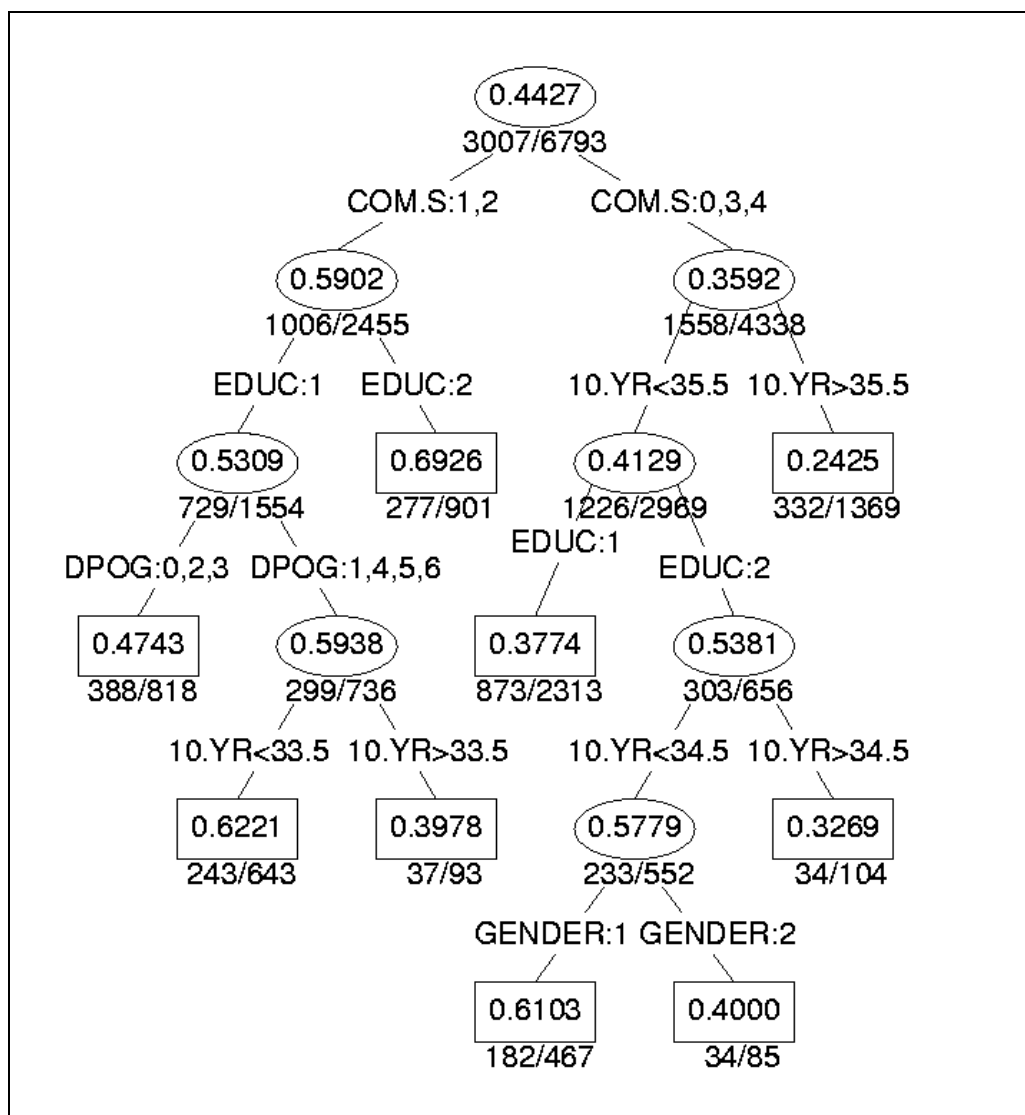Misclassification error rate: 0.3717 = 708 / 1905

Figure B.4.   Classification Tree for Pooled Data Set.

Variables actually used in tree construction:

COMMSOURCE, EDUCATION, DPOG, TENTH.YEAR.AGE, GENDER

Number of terminal nodes: 9

Residual mean deviance: 1.279 = 8675 / 6784

Misclassification error rate: 0.3533 = 2400 / 6793

LIST OF REFERENCES

Army Regulation 614-100, "Officer Assignment, Policies, Details, and Transfers", [http://www.usapa.army.mil/pdffiles/r614%5F100.pdf], September 2000.

Bowman, W.R. and Mehay, S.L., *Graduate Education and Employee Performance:* Evidence from Military Personnel, *Economics of Education Review*, Review 18, 20 December 1998.

Buterbaugh, T.A., *A Multivariate Analysis of the Effects of Academic Performance and Graduate Education on the Promotion of Senior U.S. Navy Officers*, Master's Thesis, Postgraduate School, Monterey, California, June 1995.

Branigan, G.A., *The Effect of Graduate Education on the Retention and Promotion of Marine Corps Officers*, Master's Thesis, Naval Postgraduate School, Monterey, California, March 2001.

Chambers, J.M and Hastie, T.J., *Statistical Models in S*, Wadsworth & Brooks/Cole Advanced Brooks & Software, 1992.

Conzen, E.L., *An Analysis of the Impact of Fully Funded Graduate Education on the Retention of Naval Officers*, Master's Thesis, Naval Postgraduate School, Monterey, California, December 1999.

Department of the Army Pamphlet 600_3, "Commissioned Officer Development and Career Management", [http://books.army.mil/cgi-bin/bookmgr/BOOKS/P600_3/CCONTENTS], 1998. (04.22.2003)

Department of Defense Directive, Number 1322.10, "Policy on Graduate Education for Military Officers'

Hamilton, L.C., Regression with Graphics, Duxbury Press, 1992.

Hosmer, D.W. and Lemeshow, S., *Applied Logistic Regression*, Wiley-Interscience Publication, 1989.

National Defence Research Institute, Defense Manpower Research Center, *Future Career Management Systems for U.S. Military Officers*, RAND, 1994.

National Academy of Sciences, Volume 4, p.39, "Education and Training",
[http://www.nap.edu/html/tech_21st/hr2.htm], 1997. (04.22.2003)

Naval Postgraduate School General Catalog, 2001

Occupational Data Base, "Old to New Occupational Code Conversion Table" [http://www.odb.asmr.com], November 2000. (04.22.2003)

Officer Personnel Management System XXI Task Force, "Officer Personnel Management System XXI Study", 1997.

S-PLUS 2000 User's Guide, Data Analysis Products Division, MathSoft, Seattle, Washington.

INITIAL DISTRIBUTION LIST

1. Dudley Knox Library
    Naval Postgraduate School
    Monterey, CA

2. Professor Samuel E. Buttrey
    Department of Operations Research
    Naval Postgraduate School
    Monterey, CA

3. Professor Robert A. Koyak
    Department of Operations Research
    Naval Postgraduate School
    Monterey, CA

4. Hakan Kabalar
    Kara Kuvvetleri Komutanligi
    Lojistik Baskanligi
    Bakanliklar, Ankara, TK